Segmentation Style Discovery

Kumar Abhishek[1], Jeremy Kawahara[2], Ghassan Hamarneh[1]

[1]Medical Image Analysis Lab, Simon Fraser University, Canada; [2]AIP Labs, Hungary

Variability in Medical Image Segmentation

Annotators' preferences and skill levels

Ambiguous object boundaries

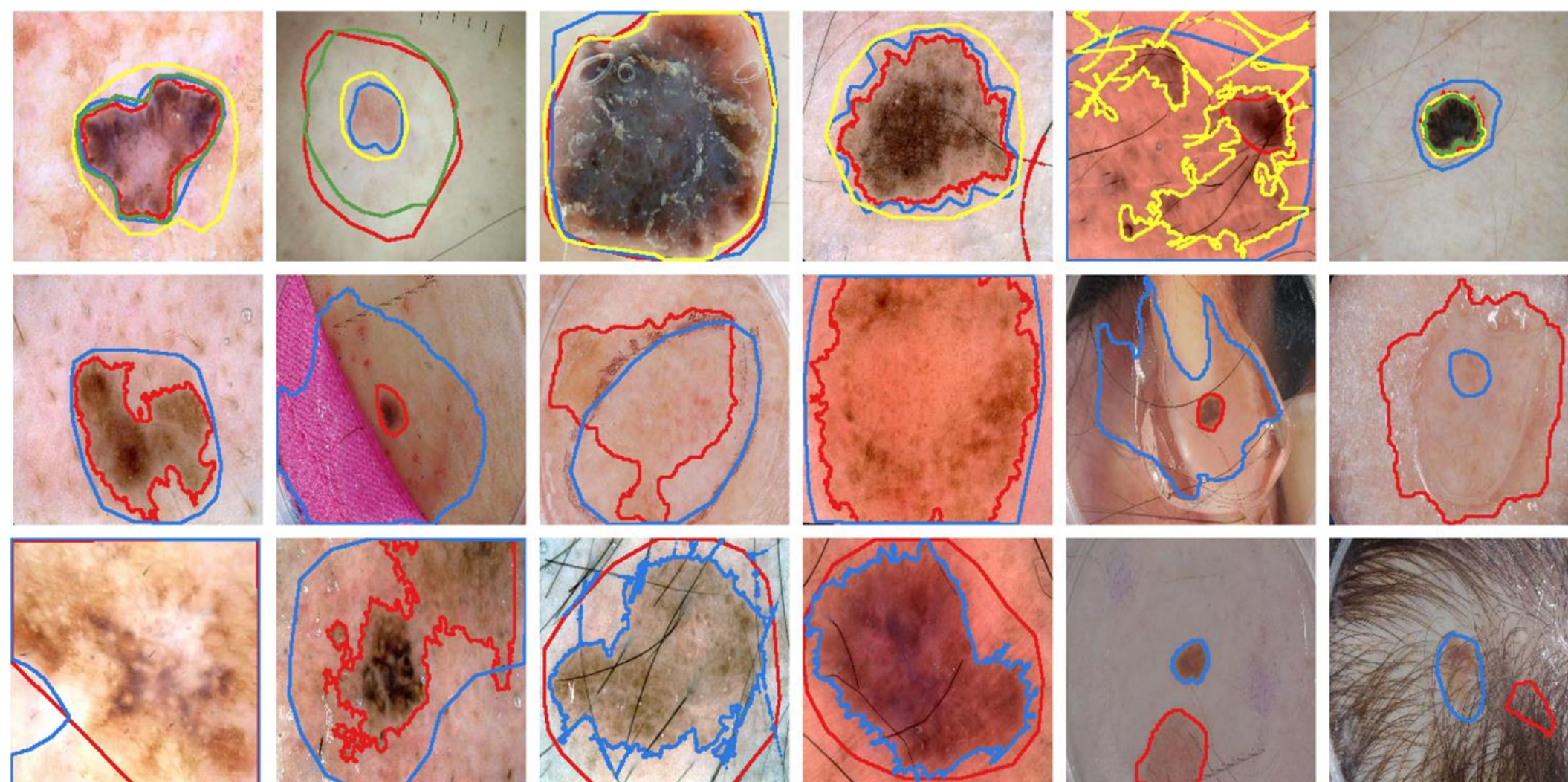Segmentation criteria and tools
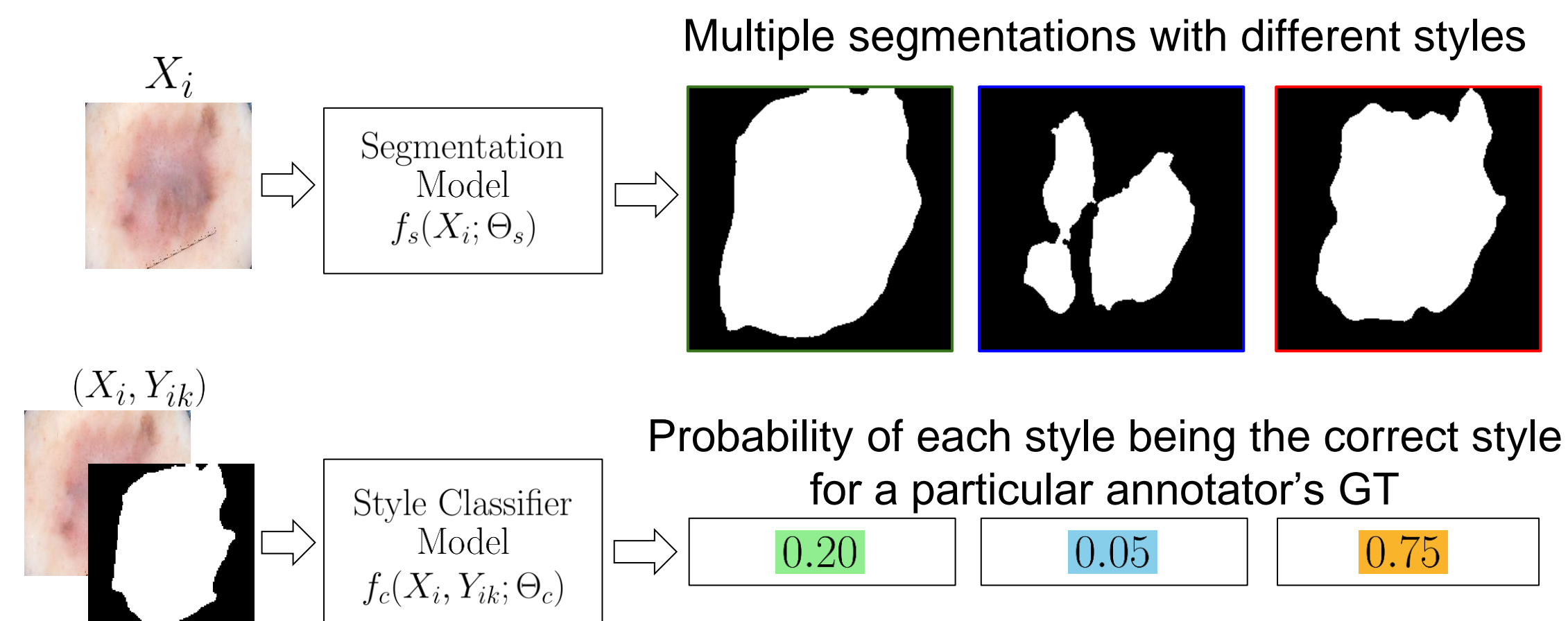
Latent factors

Motivation

Given a **multi-annotator image segmentation dataset** with or **without the knowledge of latent factors**, can we **discover** and predict **segmentation styles** that are **plausible, diverse**, and **semantically consistent** across all images?

Skin Lesion Segmentation Variability

**International Skin Imaging Collaboration (ISIC) Archive:** 2,261 images with > 1 "ground truth" (GT) segmentation mask.

StyleSeg: Method Overview

X_i

Segmentation Model $f_s(X_i; \Theta_s)$

Multiple segmentations with different styles

(X_i, Y_{ik})

Style Classifier Model $f_c(X_i, Y_{ik}; \Theta_c)$

Probability of each style being the correct style for a particular annotator's GT
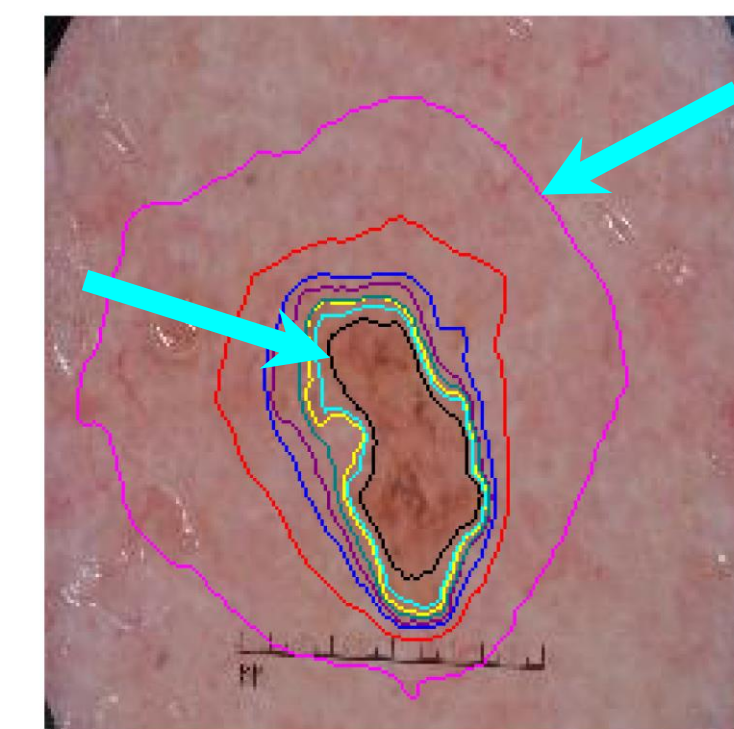
| 0.20 | 0.05 | 0.75 |

$$\mathcal{L} = \mathcal{L}_{\text{segmentation}} + \mathcal{L}_{\text{plausibility}} + \mathcal{L}_{\text{StyleClassifier}}$$
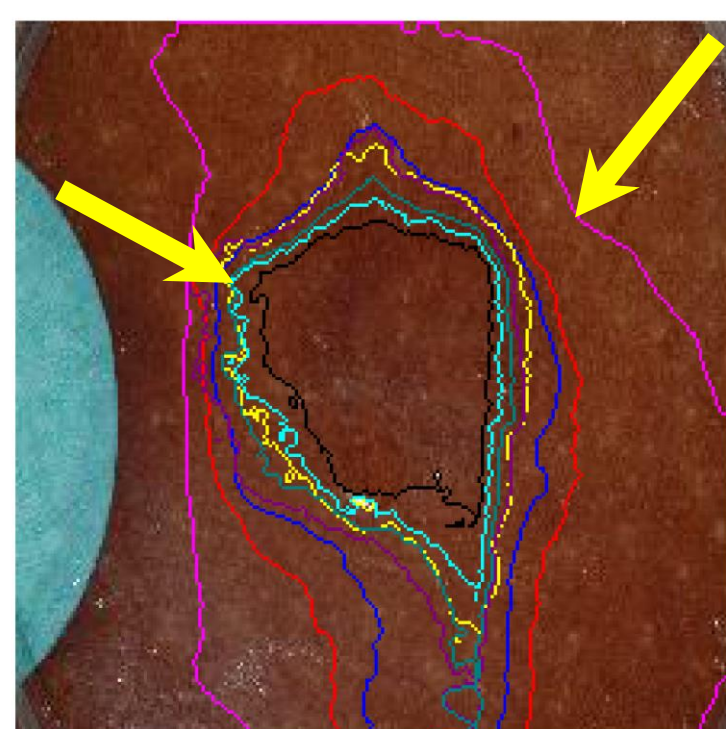
StyleSeg: Results

StyleSeg **outperforms** several **single- and multi-segmentation** prediction methods on **4 skin lesion segmentation datasets**: ISIC Archive-Test, DermoFit, PH2, and SCD.
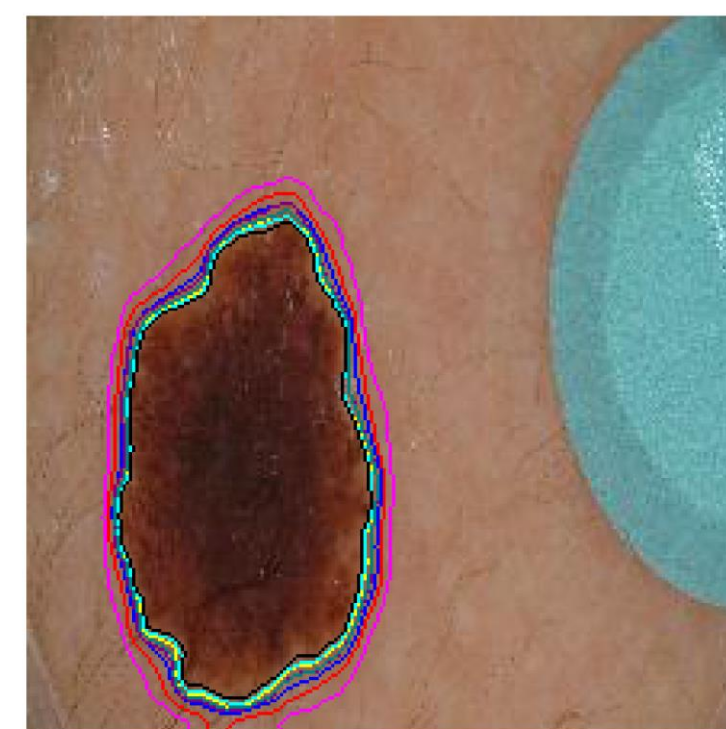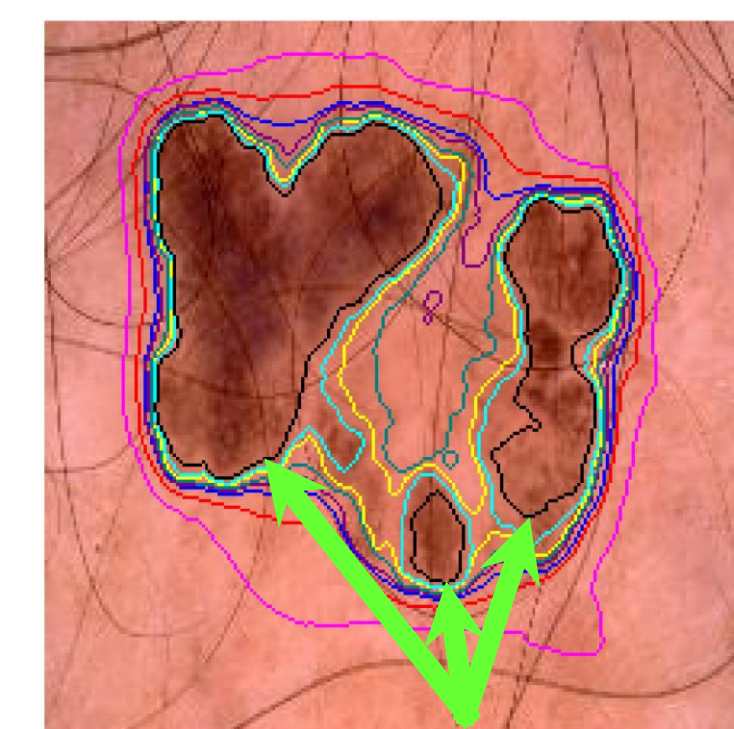

High-contrast lesion has **high agreement** across styles
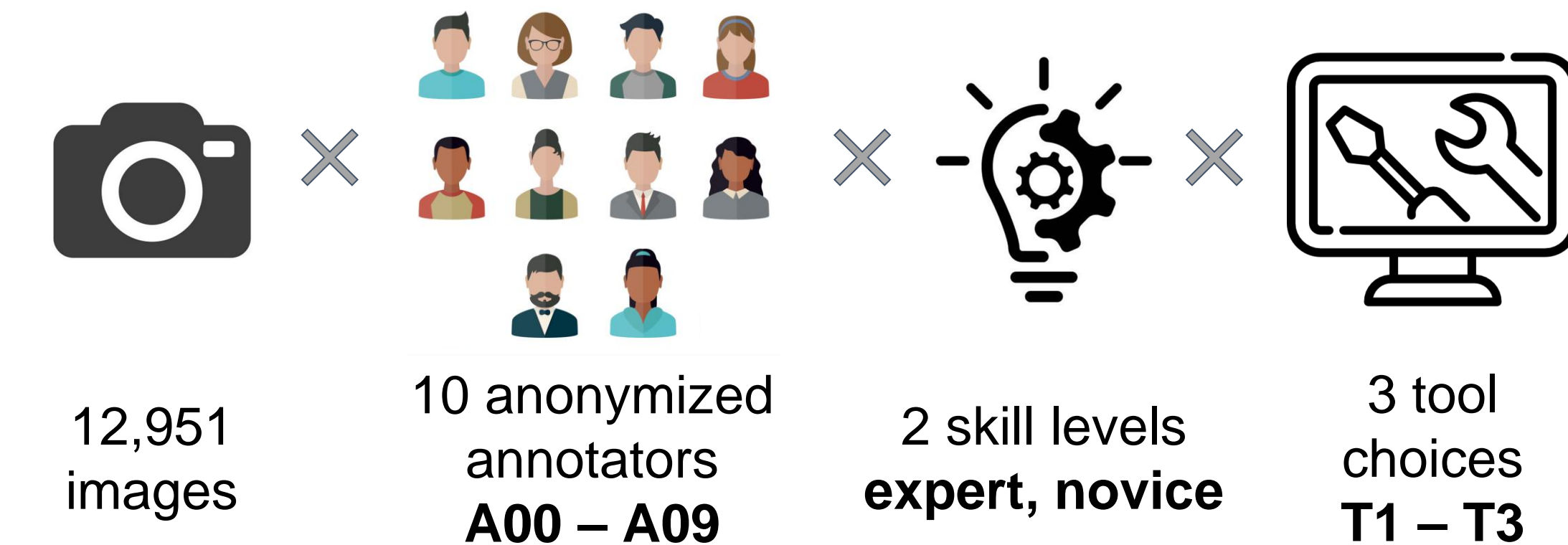
Instances of **under-** and **over-**segmentation
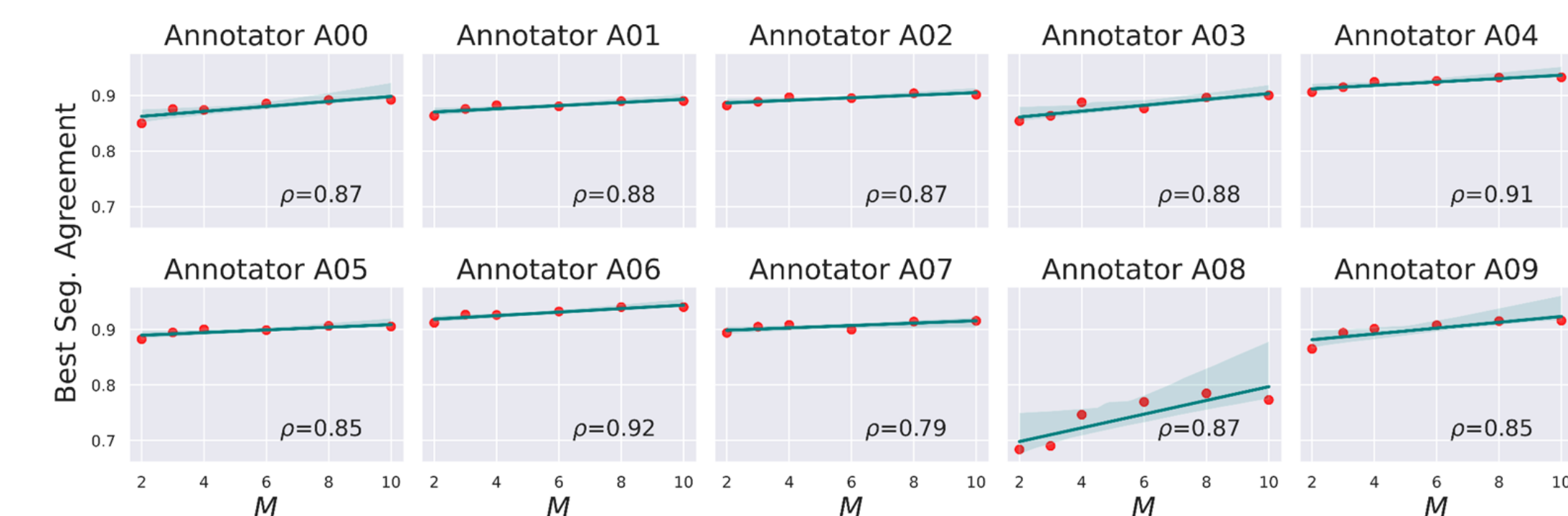
Different **boundary jaggedness** across segmentations

Ambiguous boundary causes segmentation masks to **split**

A New Dataset: ISIC-MultiAnnot

The **largest** multi-annotator skin lesion segmentation dataset.

12,951 images

10 anonymized annotators **A00 – A09**

2 skill levels **expert, novice**

3 tool choices **T1 – T3**

13,555 image-mask pairs
27 unique annotator preferences

Key Takeaways

- StyleSeg **discovers** and learns to predict segmentation styles that are **plausible, diverse**, and **semantically consistent** across all images.

- Even as the number of styles modeled increases, the styles exhibit **diversity without compromising plausibility**.

- **Personalized segmentation** (each user can choose their own style) outperforms one-size-fits-all approaches.

- ISIC-MultiAnnot: the **largest public multi-annotator** skin lesion segmentation dataset.